DOCUMENT RESUME

ED 323 537 CS 212 491

AUTHOR Haswell, Richard H.
TITLE Change in Undergraduate and Post-Graduate Writing Performance (Part 2): Problems in Interpretation
PUB DATE 89
NOTE 37p.; For part 1 of this study, see ED 269 780.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Age Differences; *Factor Analysis; Graduate Students; Higher Education; Longitudinal Studies; *Measurement Techniques; Research Problems; Skill Development; Undergraduate Students; *Writing Improvement; *Writing Research; *Writing Skills
IDENTIFIERS *Writing Development

ABSTRACT

This paper provides additional discussion and interpretation of the original findings of a study examining the development of writing skills across college years. The paper focuses on six interpretive problems from the mass of primary data in an attempt to clarify issues and block off false interpretive trails. The paper discusses: (1) the departures from purely random selection in forming the research groups; (2) the cross-sectional design of the study, with no control of possible cohort effects; (3) the validity of impromptu writing samples; (4) the openness of the writing prompt; (5) the high ratio of variables to subjects, with identification of 14 factors which account for 77 percent of total variance; and (6) the investigation of writing products only, while rhetorical intentions and compositional processes remain inferential. The paper concludes with a discussion of the dangers and benefits associated with empirical, multifactorial, exploratory, and developmental interpretation of writing. (39 references are attached.) (KEH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED323537

Richard H. Haswell
Department of English
Washington State University
Pullman, WA 99164—5020

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Richard H. Haswell
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Change in Undergraduate and Post-Graduate Writing Performance (Part 2): Problems in Interpretation

Sampling, Cohort Effects, Impromptu Setting, Open Topics, Principal-Components Factoring, Evaluation of Product

Part 1 of this paper, "Change in Undergraduate and Post-Graduate Writing Performance: Quantified Findings" (Haswell 1986) records an analysis of impromptu compositions written by three classes of college undergraduates and one group of post-graduates. The analysis was exploratory. Its immediate intent was to describe, empirically, differences in the writing of the four age-groups. It assumed that such differences would help hypothesize normative changes in post-adolescent writing performance within a typical college environment. Broadly the analysis was undertaken to advance and substantiate theory about the acquisition of writing skills or the development of writing competence during the college years. More narrowly its findings were to be integrated into this author's book-length investigation of ways to better college instruction in composition through a developmental approach.

That monograph is now complete (Haswell in press; see also Haswell 1988, 1989b). It is, however, only my own use of the original findings. The primary analysis—over one hundred quantified measures of one hundred and sixty essays composed under controlled circumstances—remains one of the most extensive bodies of empirical data available for

CS212491

study of undergraduate and post-graduate writing. My aim in the second part of the present paper is to do what I can to make that mass of primary data interpretable for other students interested in the connections between post-adolescent development and college-related writing performance.

Interpretation of that data is not easy or simple. There are many problems, some generated by the very extent of the analysis, some unwittingly by the presentation of it in the first part of this paper, and some by difficulties in abstracting usable information from specific experimental circumstances. But while these interpretive problems can never be discussed away (they are endemic to research), they can be discussed, and the discussion can clarify issues and block off some false interpretive trails. It can, for instance, forestall misinterpretation by bringing new and related facts to bear (see Section 1 below), or help untangle inter-related variables through second-step factoring (see Section 5). In short, the aim of this second part is a simple and direct one, to provide new information and discussion which will help other students of composition interpret and use the original findings.

I list the major problems in the order I will take them up.

Section 1: There were departures from purely random selection in forming the experimental groups.
Section 2: The design is cross-sectional, with no control of possible cohort effects.
Section 3: The nature of the writing studied was 50-minute impromptus.
Section 4: The writing prompt or stem was relatively open.
Section 5: There was a high ratio of variables to subjects.
Section 6: Only products of writing were investigated in any depth, and rhetorical intentions and compositional processes of the subjects can be only inferred.

This list of problems excludes several other interpretive issues taken up in Part 1 of the present paper (Haswell 1986): the chances that the student groups differed in verbal ability or academic inclination (pp. 6-9), in particular the possibility that the exiting freshman group differed from the other student groups in terms of motivation (pp. 9-11), the possibility that

self-selection entered into the sampling (pp. 5-6), the difficulty in constructing meaningful empirical measures of verbal products (pp. 13-17), and the presence of gender and topic effects (pp. 59-60). Only the first of these problems, the issue of possible group differences in pre-test writing ability, will be considered again here.

1. Departures from random selection in forming the research groups

Five age-groups were studied: eighteen-year-old entering freshmen (FR); eighteen-year-old freshmen exiting the beginning composition course (EF); nineteen-year-old first-semester sophomores (SO); twenty-year-old first-semester juniors (JU); thirty-year-old or older post-graduates (MA). These groups were formed by random selection. The original selection, however, was subsequently adjusted for two of them. (In Part 1, this adjustment is referred to as "matching"—a misleading term. The study did not use a "matched-pairs" design. In the selection of subjects there was no effort to find equivalent or "matching" scores in order to form the research groups, as was done for instance in Jewell, Crowley, and Rhum's unfortunate 1969 study.) Specifically, it was felt necessary to replace six subjects (of the total one hundred and sixty), altering the random selection of two of the five groups.

To judge how this departure from random selection may affect an interpretation of the findings, it will be useful to review the selection process. For the four student groups, the initial pool was all students in beginning and advanced writing courses at the start of the semester (FR, SO, JU) or, with the end-of-the-semester freshmen (EF), all students remaining in nine of the forty-two sections of the beginning course. Using a controlled random-selection method, subjects were picked from these pools to achieve as wide a representation of teacher as possible and to form the desired configuration of half male/half female of the right age for each group. Within these parameters, no subjects were eliminated unless they happened to be Honors students, exemptees from freshman composition, ESL writers, or students not making regular academic progress. An ANOVA was run on the resulting sample for the four groups. The criterion variable was a measure of pre-test verbal ability, namely the Verbal

Composite score of the Washington Pre-collegiate Examination (see Part 1, p. 6). For this initial, purely random sampling, there were no significant differences for main effects of group or gender or for their interactions (Main Effects: df 3; MS 96.42; F 1.17).

Nonetheless, a comparison of the distribution of Verbal Composite scores for these four student groups showed a disconcerting anomaly for the entering freshmen (FR) at the top of the distribution: the FR group had only one score in the 60's whereas the other groups had at least five. It was then realized for a near certainty that the gap had occurred because the FR population sampled was different in one crucial way from the SO, JU, and EF populations. At the time the writing was elicited, during the second meeting of class, many entering freshmen had just taken an examination to exempt themselves from freshman composition, and either were waiting to hear the results of the examination before attending class (if they passed they would drop the course) or sanguinely had not enrolled in the course at all, knowing that they could always take it the following spring semester if they did not pass the exemption test. Sixteen percent of the entering freshman class had been invited to take the exemption examination that semester, the very portion of the entering freshman class who had earned a score of 60 or better on the Verbal Composite measure of the Washington Pre-college Examination. What we realized, too late, was that although students who passed the exemption examination had been eliminated from all research groups, many verbally adept students who had been invited but had failed the examination (twelve percent of the freshman class) were present in the SO, JU, and EF populations but absent from the FR population.

To adjust for that inequality of populations, it was decided to add five subjects with Verbal Composite scores in the 60's (drawn from the original pool) and to eliminate five randomly selected students from the group. The decision was based on the distribution and variance of Verbal Composite scores reported by the university's Testing Services for the entering, high-school-direct freshman class as a whole.

One further adjustment—of a more dubious nature—was made, to the junior group (JU). One subject was randomly eliminated from the sample and replaced with another subject in order to provide a writer scoring below the first standard deviation of the Verbal Composite range. This

was done because a study of "basic" writers and development was projected. (Part 1 incorrectly reports two essays changed in the JU group, but one of these subjects was replaced because it was discovered, on typing up the essays for analysis, that the handwriting of the data sheet attached to the essay was different from that of the essay. It seemed possible that another student's data sheet had been stapled to the essay. So another essay was drawn at random from the original pool.)

All told, the adjustment made to the initial randomly selected sample is small in size. Only six individuals were eliminated out of a sample of one hundred and sixty. The initial sample for three of the research groups remained intact (SO, EF, and MA), and another group, JU, had only one individual substituted. The major alteration occurred in the FR group, where the change should bring the sampling closer to the true population, and the effect of the change should be not to hide a regression toward the true mean (as may occur in "matched-pairs" research designs) but just the opposite, to adjust the sample closer to the true mean. The best candidate for inequality in the groups probably lies in an over-representation of better writers in the FR group, which in fact would tend to understate the main implication of the subsequent findings, that students do mature in their writing during the two years of college.

I would hazard that the main problem in interpreting this particular sampling lies less in accounting for a possible inequality among the groups and more in answering the question of which population the groups represent. They do not represent student classes at large (freshman, sophomore, etc.), nor the typical group of students a beginning or advanced composition teacher faces in the classroom. They did not tap, for instance, freshmen who were not eighteen years of age or, for another instance, juniors who were Honors students or second-language writers of English. Roughly it can be said that the student groups represent a kind of scholastic middle-of-the-road beau-idéal, a "regular progress" student neither so talented nor dedicated to make Honors college or to be exempted from freshman writing, nor so side-tracked to be delayed in entering college or unable to manage full credit because of extra-curricular work or academic problems.

It should also be noted that the main goal of the study was to explore developmental-learning sequences, so that of equal importance with equalizing initial verbal ability was regularizing chronological age and

academic experience. For that purpose, the method of sampling seems forthright, although it does not lack further problems in interpretation.

2. Cohort effects

One of these problems emerges from the cross-sectional design of the study. The study does not follow the changes in writing performance of a group of students over the academic years, as would a longitudinal design, but rather samples different-age students at one point in time. This is not the place to debate the comparative fitness of the two designs (for that, see Rest), although it is tempting to note that while the dangers of the cross-sectional are well known, the presently favored longitudinal design has it own problems in recording writing change, especially the problem of controlling the effect of repeated testing and the severe trouble with attrition of subjects enrolled in institutions of higher education. But it is worth discussing one crucial problem with the present cross-sectional design, namely its inability to control for cohort effects. How can one tell whether a change in writing performance from one age group to another—a change one would like to attribute to an interaction and learning and development—is not instead due to the fact that the two groups belong to two different historical environments, two different cohorts?

The presence of cohort effects in writing is highly speculative, supported almost entirely by intuition. The tiny handful of empirical studies that seem to have detected cohort differences lack control of population and writing prompt (e.g., Cannon, Gaies, Sloan). With the present study, the bottom line is that any one of the writing-performance differences between groups may be due to historical rather than developmental causes. But it seems doubtful that many of the changes are cohort effects, for two main reasons. One is simply that a span of three years, the longest birth-year gap possible between any two of the student subjects, will not likely record much historical shift in writing style or rhetorical purpose. Basic ways of writing show only the most glacial movement through time, despite teacher laments of continuing deterioration in entering student writing. (Both state-wide and nation-wide, SAT scores of entering freshmen remained

stable during the years the three student cohorts of this study matriculated. So did the Washington Pre-collegiate Examination verbal scores.)

The second doubt looks at the kind of writing traits this study measured. They seem to be traits little sensitive to shifts in historical wind. Of use here is a distinction made by William Perry, Jr., in his review of the debate over the stability of cognitive styles through time. He distinguishes between "cognitive style" and "cognitive strategy." Cognitive style he defines as "the relatively stable, preferred configuration of tactics that a person tends to employ somewhat inflexibly in a wide range of environmental negotiations." Cognitive strategy is "a configuration of tactics chosen or constructed from an array of available alternatives to address a particular kind of environmental negotiation" (pp. 106-107). With this distinction Perry proposes two different ways that humans "negotiate" given tasks, one way using tactics that are more general, stable, and automatic, the other using tactics that are more goal-specific, malleable, and conscious.

Now the writing competencies investigated in this study tend to be of the first kind. In general, the study measures features, such as density of cohesion or breadth of vocabulary, that reflect stylistic propensities underlying more surface maneuverings, such as opening with a shocker or defining technical language for a non-technical reader. This does not mean the study looks only at tactics that cannot be changed or tend to resist learning or personal development. Perry argues that while some cognitive styles remain relatively fixed, other styles may be acquired both gradually and rapidly. But it does mean that in tapping deep-seated stylist traits rather than surface strategies, the researcher is much less apt to have measured any cohort effects. The public figures that writers allude to (current pop stars, for example) may change according to the historical year, but not likely the rate at which writers include such allusions at all (see Measure 18) or the proportion of them from pop culture (Measure 19). (The numbered measures mentioned here and elsewhere are described in Part 1, pp. 17-37). Yet both of these writing traits might well develop with a student during two years of learning in college. The group differences most attributable to historical rather than personal development (assuming that the two can even be distinguished; see Riegel) are those showing no change during the student years and a change with the post-graduates: rate

of "level-one" or most specific nominals (Measure 13), paragraph linkage (Measure 24), use of cohesive ties of reference (Measure 31), coordination of noun phrases (Measure 59), t-unit variance (Measure 87), and correctness of predication and parallelism (Measures 95 and 97). Historically, a decade gap or greater separates the oldest student and youngest post-graduate, providing the best opportunity for the presence of cohort differences. Such a presence could be located by one study only with a research design combining cross-sectional and longitudinal measurement (e.g., Schaie).

3. The validity of impromptu writing

The nature of the writing elicited from the subjects—50 minutes long, unrehearsed, impromptu—raises the issue of validity. Did not the circumstances give writers little time to show their best abilities and, worse, even preclude some essential tactics of writing, such as careful logic and complex organization? Undeniably, the study alone cannot argue that its findings will necessarily recur in polished prose either of students or matured writers. But how likely will it be for the findings to recur? How much of them records tactics which writers sensitive to a new kind of "environmental negotiation" would alter were they given more time to write, and how much records a style which "somewhat inflexibly" remains no matter how much writers are harassed or hurried?

I hazard that a similar study, using less spontaneous writing, would find differences in degree but would parallel the findings here in most essential ways. The few researchers who have broken down both kinds of writing analytically, as does Hilgers in his study of the effect of free-writing instruction, generally find the same relative results in impromptu as in multi-drafted essays, with the latter only somewhat more discriminating statistically. English teachers, in fact, may tend to exaggerate the qualitative difference between the two. Hartvigsen elicited both in-class and out-of-class essays from throughout a semester, typed and with topics rotated. Unsuspecting evaluators, college teachers, rated the in-class essays of half of the students better than or the same as their out-of-class essays.

The classroom essays were about seventy-five words shorter, which may account for most of the six percent edge in holistic score held by the dormroom essays. One kind of trait the present study relies heavily on, syntax, showed no significant differences (cf. Swan's comparable finding).

Hartvigsen did not study other possible causes for the meagre superiority of the out-of-class writing, but studies of the way students revise suggest logic and organization as unlikely candidates. Faigley and Witte gave their subjects ample time, days, to think out, write, and re-write an essay, and in the end remedial freshmen devoted eighty-eight percent of their revisions, and students in an upper-division elective expository writing class seventy-six percent of theirs, to changes that did not alter the summary meaning of their first, in-class draft. As for organization, only one percent of the remedial changes and eight percent of the upper-class changes would alter a summary of the text (and even this pittance may exaggerate, since each sentence involved in a change was counted as a separate revision). These results should come as little surprise, remembering studies like those by Pianko and Cannon that found students going about the writing task little differently in and out of the classroom: a few minutes arranging a topic, writing a first draft straight ahead in a linear manner, then reading it through once to make surface changes. But it may come as a surprise to find that the older professional writers in Faigley and Witte's study performed much the same as did the students, in fact made fewer total revisions than even the remedial freshmen. Possibly the notion of the good writer patiently crafting day after day as a sculptor in marble may stand as one of the more treacherous idols of the English teacher.

Some developmentalists argue that true development will be elicited only when subjects are pressed by more advanced mentors to achieve their "level of potential development" (Vygotsky 86) or their "optimum level" performance (Fischer and Pipp). But how would one control for peer or mentor pressure? Originally, the study chose to use in-class writing in order to assure all writers the same fair shake, to equalize opportunities. Giving writers interim time to work over a topic might easily have helped one group over another, as where entering freshmen might have had fewer acquaintances they would have been willing to milk for ideas. The loss of validity to keep test control is a loss, of course, that all experimental research has to bear. Ideally, both kinds of writing would have been

elicited. Then comparison would have added some light to the obscurity that still prevails over the stability of writing traits relative to composing time.

4. The openness of the writing prompt.

The subject matter of the two topics themselves encouraged the writer to respond in an unforced or open manner. Although the topics specify an audience and a purpose (see Part 1, p. 4), they still define the rhetorical task rather loosely. For interpretation of the resulting compositions, the problem is that such looseness in assigned task may allow writers to assume a wide range of rhetorical intentions. In turn that variety may affect the measured traits in ways difficult if not impossible to distinguish from developmental-learning effects (Faigley 1980). Crowhurst and Piché and many subsequent researchers, for instance, have found that rhetorical aim and mode adopted by a writer have significant effect upon syntactic measures such as t-unit length. If the post-graduate writers in the present study responded to the topics with less of a narrative treatment than did the students, for example, then that alone might have caused the significant increase in t-unit size with age (Measure 54).

The traditional developmentalist answer to this quandary is that <u>how</u> subjects select rhetorical mode and purpose is itself developmental. Among researchers in composition and administrators of assessment, the validity of loose topic is still up for debate (see Brossell's summary), but among developmentalists openness of prompt has long been the most common methodology. A narrow prompt may elicit a less developed response than the subject is actually capable of. Following the practice established by Freud and Piaget, researchers such as Perry, Jane Loevinger, Rita Weathersby, Carol Gilligan, and Mary Belinky and her colleagues have set the most open-ended queries or problems to let the deepest and most sophisticated emotional and cognitive structures emerge to handle them. Similarly, this study preferred a wide net, trusting that the broad topic would attract different responses and that the broad spectrum of measures would record them..

The study also assumed that such a range of measures would help interpret these differences. For instance, the measure of rhetorical mode showed no difference among groups in the amount of essay space taken up with modes that go beyond factual reporting, such as making value judgments and recommending action (Measure 3). Yet the older writers preferred logical organizations of inference and problem-solving, suggesting that they may have found ways to direct their essay toward a conclusion of value or policy without devoting too much space to the actual expression of it (Measure 4 and Table 1 of Part 1). For another instance, the older writers more often make the first-person subject of their sentences and presumably as locus of their sentence topic (Measure 45), but that does not simplify their syntax (Measures 53-55, 67-68). For a third instance, the older writers organized their full essays in a narrative manner (through patterns of development or causation) considerably less often than did the students, yet on the local level they organized sections of their paper as often through such patterns (see Tables 1 and 2), suggesting that they relied on the narrative mode as much as did students but subsumed it under more complex, overarching logical purposes. All told, I think that the benefits of the relative openness of the topics outweigh the hindrances, but that openness requires interpreting the results with interactions among measures always in mind.

5. The high ratio of variables to subjects: a factor analysis

Unfortunately, interactions do not have to multiply very much before they transgress limitations, limitations statistically in the degree of sample bias tolerable and mentally in the number of relationships a person can hold in mind and make interpretive sense of. Fortunately, ways to handle the statistical problem help deal with the conceptual one. The standard way in human-science research, of course, is through factor analysis, although that way has not been much used by composition researchers, who usually deal with a limited number of empirically measured variables (exceptions are Diederich et al., Zaharias and Mertz, and Linn and Piché). Given the

application of such a large battery of variables in the present study, factor analysis seems urgently called for.

This study applied eighty-three pre-set variables to one hundred and six subjects—at about a 1:2 ratio. The statistical rule-of-thumb is that in multivariate analysis the variable-subject ratio ought to be at a minimum 1:10 and preferably closer to 1:20. The technical reasons are complex. But in terms of interpretation of findings, one danger of a low subject-variable ratio is paramount. Any trend in the sample involving more than one variable may be spuriously inflated by high correlations between variables. This study, for instance, found the sophomore or junior groups differing from the freshman group in twenty-one of the eighty-three measures, and found the post-graduates differing from the undergraduate groups in forty-four of the eighty-three. The two findings constitute the essential support for the presence of major maturing change in writing performance taking place during and after college. Although the ANOVAs run individually on these measures (each incorporating an its own adjustment for sample size) indicate statistical confidence in individual measures, that confidence does not extend to the general conclusion because nothing is known about the inter-relationships among the variables. The twenty-one undergraduate changes may be so highly inter-correlated that they express much the same factor (multicollinearity), and what looks like a massive and broad array of changes may boil down to only one kind of change. The same may even be true of the forty-four working-world changes. The fact that these forty-four changes touch on all seven of the categories used to organize the eighty-three pre-set variables (Overall quality, Ideas, Support, Organization, Diction, Syntax, and Mechanics) is an argument in favor of the breadth of the changes but statistically not a very convincing argument, since those categories were intuitive and apriori, based on no quantitative evidence that they represent statistically independent factors.

Statistically, the obvious way to settle the argument is to run a principal-components analysis of the eighty-three variables. This would create factors perfectly non-correlative to each other and would show the relationship of individual variables to these factors. Such analysis was not done with the study originally, but it has been done since, and I will set forth the results here. There are several benefits to this post-facto principal-components analysis. It will aid interpretation by unscrambling a

tangled host of variables and identifying major factors underlying the traits studied, it will reduce variables down to an acceptable variable-subject ratio, it will recommend an optimal set of variables for use in further studies of writing development, and it will test the hypothesis that the undergraduate and post-graduate change in writing performance is broadly based and not restricted to a few factors.

Three prinicpal-components analyses were run on the measures. Only the first will be reported in full here, since the second two were run primarily as a check upon the first. All analyses were formatted by SPSS^X (1986), applying the principal components analysis of Harman (1976). The factoring was not rotated because the data set did not meet Thurstone's criteria for simple structure. The first run used eighty variables of the full eighty-three. The three variables eliminated from the analysis were Holistic evaluation (Measure 1), since it was dependent on the other traits; Essay length in words (Measure 2), since word-length was the standard denominator used to compute rates for most of the other measures and consequently this particular measure would generate the primary index effect; and Number of examples (Measure 12), since it was a measurement that did not adjust for length of essay. The second run used eighty-two variables, including Essay length in words and Holistic evaluation in spite of their statistical dubiousness, in order to see with which factors those two important variables might correlate highly. The third run used seventy-four variables, eliminating not only the three variables ejected from the first run but six others, all "basket" variables that collected the analysis of other measures under one sum: Logical indicators (Measure 26), Cohesive ties (Measure 28), Free modification (Measure 71), and so on. The intent was to control for inflated whole/part effects. As it turned out, the resulting factoring was essentially the same in all three runs, although differences in factor loadings did aid in subsequent analysis of the results. Here I will report the results of the first run, alluding to the second and third runs only where instructive.

Kaiser's rule recommends that interpretation of a catoring should consider only factors with Eigenvalues of 1.0 or better. The analysis extracted twenty-three such factors, with a cumulative percent of variance totalling 93.1 (see the scree plot in Figure 1). Of these twenty-three

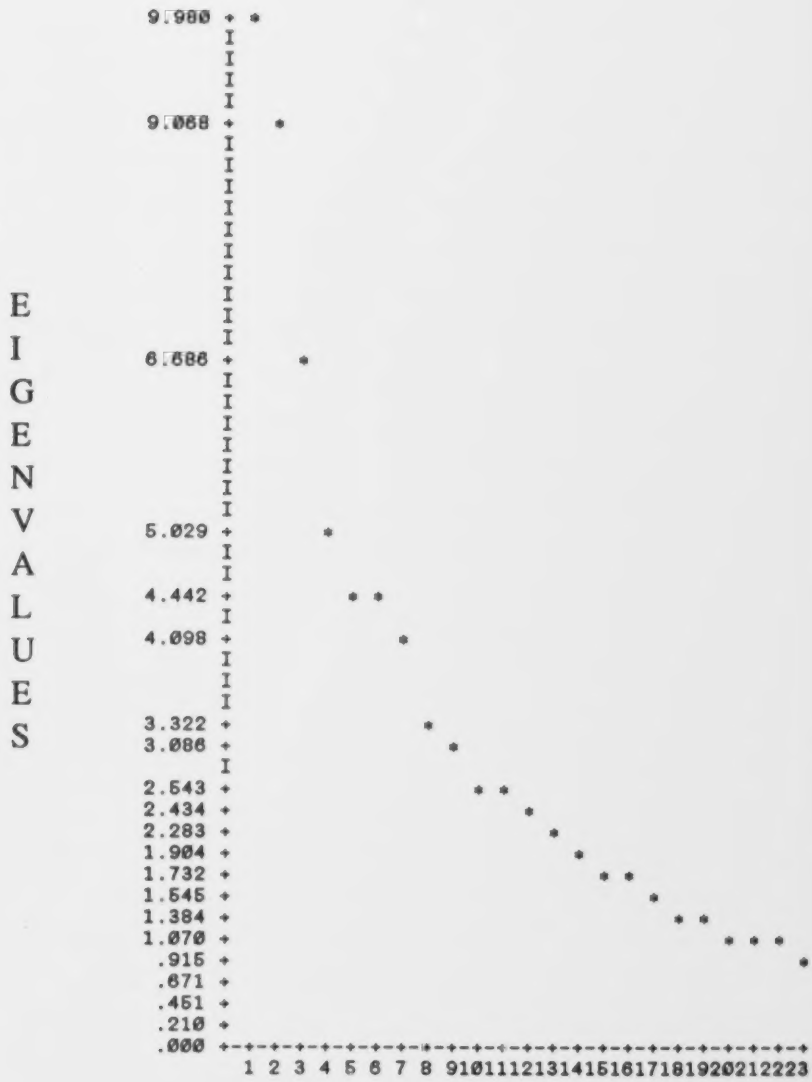

Figure 1. Scree plot of a principal-components factoring of eighty-three measurements of writing performance, resulting in twenty-three factors with Eigenvalues of 1.0 or better.

factors, nine did not contain at least one variable with an acceptable loading, of .400 or better. The remaining fourteen factors are listed below, each with its top-loading variable. A single asterisk following a variable indicates that the measure was one of the twenty-one showing statistically significant differences between student groups, and a double asterisk indicates that the measure was one of the forty-four showing a difference between the post-graduates and students (this use of asterisks will be continued in following lists).

Factor	Eigen-value	Pct. of Var.	Top-loading Variable	Measure		Loading
1	9.98	12.5	T-unit size	54	**	.954
2	9.07	11.3	Clause size	55	**	.831
3	6.69	8.4	Logical depth	8 *	**	.712
4	5.01	6.3	Relative clauses	61 *		.636
5	4.49	5.6	Logical indicators	26		.578
6	4.44	5.6	Rhetorical mode	2		-.608
7	4.10	5.1	Free modification	71	**	.564
8	3.32	4.2	Passives	42		.571
9	3.08	3.9	Level-four abstractions	6		.650
10	2.69	3.4	Synonym ties	30		.581
11	2.54	3.2	Exemplification	11 *		.550
12	2.43	3.0	Length of top logical pattern	4 *	**	.479
13	2.28	2.9	Syntactic parallelism	85 *		.415
18	1.44	1.8	Correct predication	95	**	.404

Statistically, of course, these fourteen factors are perfectly orthogonal to each other, and the top-loading variables stand together as best choice for a cluster of measures representing those factors. Reducing the eighty-three variables to these fourteen would offer a tolerable variable-subject ratio (about 1:11). The fourteen factors account for seveny-seven percent of total variance.

The most important implication of this factoring is that it still supports the initial finding of broad change during and after the college years. The undergraduates show change in thirty-six percent of the fourteen factors,

the post-graduates forty-two percent. So at first reading, this factoring does not change the interpretation of undergraduate writing development encouraged by the full spectrum of measures.

Any principal-components analysis, however, particularly of this size, needs some interpretation itself. The top-loading variables reflect their respective factors to a greater or less degree—generally less—and what quality or competency of writing these factors reflect cannot be told very well from just one representative variable. Scrutiny of the other high-loading variables for each factor is needed to help characterize that factor. As we will see, with this factoring of principal components such a scrutiny, factor by factor, not only assists in unscrambling the full set of variables, it also further supports and clarifies the picture of substantial change in writing performance during college. I take up the factors in order, listing the variables with loadings of .400 or better, and hazard a brief characterization of the factor.

Factor 1: Variable	Measure	Loading
T-unit size	54 **	.954
Sentence size	53 * **	.941
Independent clauses	66	-.724
Clause/t-unit ratio	56	.721
Mono-clauses	65	-.662
T-unit variance	87 **	.659
Second-person plural	46 * **	-.621
Dependent sentence openers	86	.592
Post-nominal modification	69 * **	.537
Punctuation of main clauses	99	.514
Pre/post-nominal mod. ratio	70 * **	.503
Nominal complexity	68 * **	.502
Prepositional strings	80 **	.497
That clauses	62	.492
Coordination of t-units	57	-.483
Free modification	71 **	.481
Nominal modification	67 * **	.476
Subordinate clauses	60	.455

Factor 1 (cont): Variable	Measure		Loading
Initial adverbs	76		.444
Predicate coordination	58		-.450
Correct parallelism	97	**	.428
Passives	42		.410

This factor, Factor 2, and Factor 7 partial out the mass of syntactic measures. Roughly, Factor 1 represents what might be called multi-clause unit-span, the breadth of interdependent ideas a writer is able or willing to encompass within a unit that writer conceives of as independent. Mere sentence length and t-unit length stand in nearly perfect correlation with the factor, allowing easy interpretation. Percent of total words put into independent clauses (Measure 66) and into single-clause t-units (Measure 65) correlate negatively with the factor because these stylistic preferences co-occur with a bent for shorter independent syntactic units: the fewer the words put in dependent structures, the shorter the independent structures of sentence and t-unit tend to be. The same reason may explain why coordination of predicates and t-units also correlate negatively. Of interest is the negative correlation of preference for the impersonal you ("You can turn strangers off with improper dress"). Perhaps the usage (which the post-graduates eschewed) lends itself to advisory and therefore terse utterances. The high and positive loading of the measure of clause/t-unit ratio helps clarify the kind of syntactic breadth this factor represents, because that variable increases as more and more dependent clauses are attached to t-units. The positive loadings of dependent sentence-openers, modifiers, prepositional phrases, free modifiers, adverbial phrases, and subordinate clauses are self-explanatory. Finally, the high correlation of t-unit variance with the factor suggests that such variance—usually taken as a sign of highly mature writing—is achieved with this kind of multi-clause syntactic span. Typically, writers achieve high variance by generating exceptionally long t-units that vary with a few short ones.

Assessing the top-loading variables as a whole, it can be seen that both undergraduates and post-graduates show a pervasive developmental movement with Factor 1.

Factor 2: Variable	Measure	Loading
Clause size	55 **	.831
Nominal modification	67 * **	.675
Mono-clauses	65	.622
Clause/t-unit ratio	56	-.600
Prepositions	79 * **	.596
Pre/post-nominal mod. ratio	70 * **	.594
Long words	38 * **	.586
Dependent sentence openers	86	-.568
Subordinate clauses	60	-.558
Independent clauses	66	.544
Standard sentences	98	.535
Prepositional strings	80 **	.534
Correct spelling	92 **	.512
Pronominalization	44 **	-.507
Base clauses	64	-.468
PLOM words	37 * **	.474
Predicate coordination	58	-.467
Qualifiers	10	-.465
Correct contractions	94	.424
Initial adverbs	76	-.432
Punctuation of main clauses	99	.427
Top-level logical pattern	4 * **	.416
Adverbs	75 *	-.414

Factor 2 seems to represent sub-clausal syntactic span, the ability to conceive and utilize more and more sizeable, semantically bound nodes of ideas. The factor seems to verify John C. Mellon's argument that one major feature of verbal maturity is a growing attraction toward "restrictive structure of dominant NPs [noun phrases]" (p. 17). Size of clause stands, of course, as the central syntactic form measuring this attraction, and nominal modification—as Mellon saw—the main syntactic tool to satisfy it. It is worth noting, however, that the variable of "base" clauses (Measure 64) loads negatively with this factor, whereas "mono-clauses" loads positively. The base clause here measures the portion of total words

occurring in the main clauses of multi-clause t-units; mono-clause measures the words in single-clause t-units. The respective loadings mean that with these writers expansion of dominant NPs is associated with single-clause t-units but not with more complex ones—support for the theory of Francis Christensen (which Mellon opposed) that syntactic maturity expresses itself not with expansion of the base clauses but elsewhere. We will see Christensen's theory supported by Factors 4 and 7, suggesting that both he and Mellon were partly right. Generally, the other loadings with Factor 2 offer a forthright reading, with nominal modification and prepositional phrases (a major vehicle of bound modification) correlating positively, and dependent sentence openers, subordinate clauses, predicate coordination, qualifiers (usually non-restrictive syntactically), and adverbial constructions correlating negatively. Somewhat unexpectedly, and of diagnostic interest, is the positive correlation of the two measures of vocabulary sophistication: letter-length of words and infrequently used words (PLOM). It seems that, syntactically, enhancement of vocabulary is associated with growth at the subclausal and restrictive level, not so much at the t-unit or non-restrictive level. To be sure, vocabulary attends to other rhetorical needs (see Factor 10). But the association of bound expansion within the clause with facility in word choice may connect with three other positive loadings here implicated with a sureness of language use: correct or conventional punctuation of sentences (avoidance of "sentence fragments"), spelling, and use of contractions.

Factor 2 encompasses several important maturing changes during the undergraduate years (in nominal modification, prepositions, vocabulary sophistication), and compared with Factor 1 marks an even more pervasive maturing configuration in the post-graduate writing.

Factor 3: Variable	Measure		Loading
Logical depth	8	* **	.712
Identical cohesive ties	29		-.600
Correct syntactic parallelism	97	**	.580
Elegant variation	40		.571
Level-one generalizations	13	**	.530
Size of 2nd-level logical pattern	7	**	.530

Factor 3 (cont): Variable	Measure		Loading
Cohesive ties	28	**	-.509
Size of top-level logical pattern	5	**	.475
Unique events	17		.470
Pre/post-nominal modif. ratio	70	* **	-.461
Length of conclusion	23		.458
Allusions	18		.444
Paragraph linkage	25	**	-.439
Correct spelling	92	**	.436
Conclusion	22	*	.413

Factor 3 embraces what is commonly called elaboration of ideas. The variables of logical organization here (Measures 8, 7, and 5) describe the depth or inner expansion of logical points rather than the breadth or extent of connected points, and there is an obvious association with synonym preference (over identical cohesive ties), most specific or "level-one" sentence topics, descriptions of unique events, and allusions to well-known persons, places, and events. It is important to see that the factor describes the elaboration of ideas rather than their specification (contrast Factor 4) or their support (contrast Factor 11). The association with the size and rhetorical quality of essay conclusions is insightful. The higher rated conclusions tended to expand on ideas presented in the body of the essay, not to summarize them. This distinction is iterated with the negative loadings of the variables counting cohesive ties (Measures 29, 28, and 25). Excessive use of identical ties or transitional phrases, as opposed to synonymous ties which add new information, associates with writing that does not move, does not elaborate ideas (an argument intuitively made from the cohesion measures alone in Haswell 1989b). Finally, the association of this factor with correct parallelism and correct spelling remains somewhat puzzling. The first does suggest writers who are able to elaborate syntactically, in a logically exact manner—new ideas evolving in a parallel manner from old ones.

From the developmental perspective, Factor 3 makes a clear statement, that the feature is distinctly a mature trait, with student writers showing only piecemeal movement toward it. It should be noted that in the second

principal-component extraction run, where the variable of holistic rating was retained, that measure fell under this factor (and under none other), loading positively at .513.

Factor 4: Variable	Measure	Loading
Relative clauses	61 *	.636
Reference cohesive ties	29 **	.564
Final free modification	74 * **	.558
Nominal complexity	67 * **	.497
Level-one generalizations	13 **	.464
Introduction	20 * **	.460
Adverbs	75 *	-.447
Infinitives	81	-.428
Initial adverbs	76	-.414
Length of introduction	21	.404
Base clauses	64	.402

Somewhat enigmatic, Factor 4 still seems to represent what might be called local specification of old or given information. Cohesive ties of reference often connect a new utterance back to an old topic locus. To expand on that given locus, convenient tools are relative clauses, final free modification, and logical restriction of nominals through layers of modification (Measure 67). Level-one or most specific generalizations expand it semantically. On the other hand, adverbs and infinitives are more likely signal a turn to a new locus—hence their negative loadings here. The positive association with rhetorically effective introductions and with lengthy introductions is understandable. The first measure rewarded writers who predicted well the scope and purpose of the body of their essay (see Table 4, Part 1). And mere length of introductions may reflect writers accustomed to the rhetorical maneuver, applicable as well on the local level, of introducing a topic and then returning to specify it.

The developmental history of the factor is somewhat complex. While the factor apparently is actively under development during the undergraduate years, the choice of tools to express it also develops. For specification of given information, undergraduates show an increasing

preference for relative clauses, ties of reference, free modification, and nominal modification, and, in competition with the factor, for adverbial constructions. The competent working-world writers use fewer adverbial and reference ties and do not continue the advance in relative clauses. Rather they prefer free modification, nominal modification, and more specific sentence topics.

Factor 5: Variable	Measure		Loading
Logical indicators of cohesion	26		.578
Introduction	20	* **	-.512
Noun adjuncts	39	**	-.503
Non-additive logical indicators	27	* **	.467
Essay unity	13	* **	-.446
Correct punct. of main clauses	99		.434
Size of 2nd-level logical pattern	7	**	-.429
Qualifiers	10		.407

Factor 5 seems to represent a centripetal attraction inward toward contained logical coherence, the kind of writing concern that creates tightly knit and localized centers of argument. Along with a preference for explicit statements of logical transitions goes a carefulness of punctuation of independent clauses (Measure 99), logical nicety of qualification, and (in the second and third factoring runs), a preference for exemplification and a precision in expressing syntactic parallelism. It is not surprising that negative correlations with the factor are logically organized whole essays (almost all the non-unified essays had double theses, suggesting a fixation on local ideas), functional introductions (which provided a thematic ordering for the entire essay), and breadth of the largest logically organized part (Measure 7). To interpret this last, it helps to note that the second factor extraction found a negative loading of .465 for the top-level logical pattern, implying that for essay organization this factor associates with static partitional schemes and not the more progressive argumentative schemes which allow authors logical routes from one locus of reasoning to another.

Developmentally, there is strong movement here, although it does not show up in the top-loading variable. Both undergraduates and post-

graduates gravitate away from heavy use of local cohesion. The significant group differences in Measure 27, the one developing variable correlating positively with Factor 5, actually makes the same point, since the juniors and even more the post-graduates show a rate of logical transitions decreasing from the rates of the freshmen and sophomores.

Factor 6: Variable	Measure	Loading
Rhetorical mode	3	-.608
Nominal coordination	59 * **	.508
Cohesive ties	28 **	.436
Initial free modification	72	.425
Initial adverbs	76	.416
Free modification	71 **	.414
Syntactic parallelism	85 *	.405

The measure of rhetorical mode, here with a negative loading, gave a higher rate to essays that indulged in evaluation or recommendation than to essays that confined themselves to the expositional modes of definition and substantiation. Factor 6, then, seems to represent the impulse toward the descriptive establishment of static, logical boundaries. This is often the intention when nominals are coordinated, and initial non-restrictive constructions usually serve to establish place and time. It is probably relevant that the second factoring run, which eliminated the collective measure of cohesive ties (Measure 28), replaced it in this factor with the measure of identical cohesive ties. Such ties more often reflect a writer defining or elaborating within one logical site than a writer logically displacing to new locales.

There is little evidence of developmental change in this factor, although the post-graduates show a remarkable jump in nominal coordination. Their increase in free modification centers mainly on constructions in the terminal position, which usually serve a quite different purpose from definition.

Factor 7: Variable	Measure		Loading
Free modification	71	**	.564
Final free modification	74	* **	.503
Final adverbs	78	*	.478
Appositives	84	* **	.421
Adverbs	75	*	.419
Substitution cohesive ties	32		.412
Conclusion	22	*	-.407

With a nicely circumscribed set of variables, Factor 7 clearly describes non-restrictive syntactic elaboration. It supports Christensen's venerable argument that a distinct ingredient of sophisticated written discourse is the topical expansion of new ideas outside the main clause in semantically and syntactically "free" structures. The association with cohesive ties of substitution is not so curious when one thinks how often a logical progression by analogy or similarity is expressed by such ties ("so," "likewise," etc.).

In further support of Christensen, and of later researchers such as Faigley (1979) and Watson, the evidence here strongly argues that non-restrictive elaboration is under active development during and after the college years.

Factor 8: Variable	Measure		Loading
Passives	42		.571
Length of top logical pattern	4	**	.515
Length of introduction	21		-.459
First-person singular	45	* **	-.431
Standard sentences	98		-.422
Paragraph linkage	24	**	.408
Paragraph length	25		.401

Factor 8 represents fluency of ideas or, more exactly, fluency in the expression of ideas. As has been noted, total essay length in words, a variable which ought to measure such a factor fairly directly, was omitted from this principal-components run, but when it was replaced in the third

run, it appeared under this factor. It loaded only at .424, however (that third run, in fact, created a new and different factor with word-length of essay as its top-loading measure). Evidently Factor 8 in the present run represents not just brute speed in getting words down on paper, but more comprehensively an ease in flow and linkage of ideas (see the discussion of "Production" in Chapter 8, Haswell in press). To aid flow of thought, the passive usually serves a cohesive purpose, and the sentence fragment (counted negatively in Measure 98) reflects a colloquial fluidity of expression. The word length of paragraphs and of the top logical pattern obviously show a linked fluency of thought, as does success in cohesively tying one paragraph with the next. In contrast, long and sometimes over-ambitious introductions often precede short, blocked essays, and the centering of sentence topics within the writer's self (Measure 45) may inhibit fluent production of ideas. (The second extraction run added letter length of words as a negative association, suggesting that concern for long or sophistication words may also slow down fluency.)

Evidence for development of this factor is not compelling, although the post-graduate writers did produce essays that were over a hundred words longer than those of the entering freshmen (but so did the exiting freshmen). Too many variables associated in a competitive way with this factor are developing side by side—use of the first person and of more sophisticated diction on the one hand, logical elaboration and paragraph linkage on the other.

Factor 9: Variable	Measure	Loading
Level-four generalizations	16	.650
Level-two generalizations	14 * **	-.570
Cohesive substitution ties	32	-.498
Cohesive identical ties	29	.434
Top-level logical pattern	4 * **	.405

Factor 9 describes what might loosely be called the power of abstraction, an ability to collect and connect large bodies of information. "Level-four" generalizations are subjects of t-units standing at the most abstract end of a four-tier classification according to degree of logical inclusion. So

"vegetation" is level four, "willow" level two; or "physical appearance" is level four, "shirts" level two. With Factor 9 the contrast is repeated in the area of explicit cohesion, with identical ties tending to expand a given generalization and substitution ties tending to displace or move from one generalization to another.

As with Factor 8, what developmental movement appears here seems contradictory, implying little overall change in the factor during the years studied.

Factor 10: Variable	Measure	Loading
Synonym cohesive ties	30	.581
Elegant variation	40	.497
PLOM words	37 * **	.411

This factor, and the remaining ones, require conjecture more than interpretation, with few variables to work with and those few loading not strongly. Factor 10 suggests a facility or creativity with vocabulary. "PLOM" words occur infrequently in print, and the measure was one of the strongest predictors of change from freshman to sophomore and junior, and from undergraduate to post-graduate.

Factor 11: Variable	Measure	Loading
Exemplification	11 *	.550
Pronoun agreement	96	.422

The factor may touch upon explanatory support (the usual function of such explicit examples as were counted in Measure 11), or perhaps, in conjunction with a sense of exact pronoun use, upon rhetorical clarity. Since Measure 11 found a significant decline in exemplification with sophomores and juniors from freshman levels, a decline made up by the post-graduates, if there is any developmental connection here it is of a temporary regressive or quid-pro-quo relationship with other components on the move.

Factor 12: Variable	Measure	Loading
Length of top logical pattern	5 * **	.479
Adverbial participles	82	-.439

The second extraction run added length of conclusions, with a positive loading (and with a statistically significant improvement after the freshmen year). This suggests the factor represents a kind of facility or fluency with overall logical organization, since top-rated conclusions fulfilled the logical progression of the body of the essay. Developmentally, that facility shows strong change.

Factor 13: Variable	Measure	Loading
Syntactic parallelism	85 *	.415

The second and third extraction runs added complexity of the second-level logical organization, loaded negatively, suggesting that this factor may involve a stylistic concentration upon balance to the detriment of progressive thought. Again the developmental change reflects a regression, with post-graduates returning to the lower parallelism rates of the freshmen and sophomores.

Factor 18: Variable	Measure	Loading
Correct predication	95 **	.404

The second and third extraction runs added a positive association with medial adverbs and with appositives. Both are rather advanced stylistic maneuvers, suggesting that the factor may have to do with a sense of conventional published or sophisticated expression. All three measures show a marked advance with the post-graduates.

This analysis of factors with their associated variables allows us to return to the question of whether the principal-components extraction as a whole supports the claim that the eighty-three variables show evidence of

compositional maturing during the college years on many fronts rather than on just a few. The results of the analysis now look like this:

	Factor	Change from freshman to junior	Change from undergraduate to postgraduate
1	Multi-clause unit span	yes	yes
2	Subclausal syntactic span	yes	yes
3	Elaboration of ideas	in part	yes
4	Local specification of given information	yes	yes
5	Local coherence	yes	yes
6	substantiation of logical boundaries	no	no
7	Non-restrictive elaboration of new ideas	yes	yes
8	Fluency of ideas	no	no
9	Power of abstraction	no	no
10	Creativity of vocabulary	in part	in part
11	Support/clarity?	no	no
12	Overall organizational fluency?	yes	yes
13	Balanced style?	in part	no
18	Sophistication of expression?	no	yes

In final analysis, the principal-components factoring generally argues that the high ratio of variables to subjects did not spuriously inflate the evidence for a multilateral advance in writing performance during the college years. The factoring may shows that the original analysis by separate ANOVAs artificially aggrandized the amount of development active in undergraduate writing performance because the full eighty-three measures had such a preponderance of syntactic measures. But there is still ample evidence of change that can be reasonably called developmental occurring on quite a few fronts other than the syntactic.

Of course, this factoring does not claim that these particular fourteen factors are the only or main components of collegiate writing development.

They are just the main statistically non-correlated components underlying this particular array of eighty-three variables. Left unknown are the other features of writing important in post-adolescent development. Nor does the factoring insist on any ranking of variables according to their importance in understanding or pedagogically furthering that development. Because prepositional use or complexity of top-level logical patterning, so characteristic of the post-graduate writing, did not appear as separate factors does not mean that they are insignificant qualities of writing maturity. Similarly, just because variance of t-unit length appears only once, partialed away to a subordinate role in Factor 1 beneath sentence length, does not mean that rhetorically such variance is not a more useful feature for teachers to emphasize. Interpretation for the pedagogical use of these eighty-three measures remains.

Still, the factoring provides a certain measure of help with that interpretation. It helps, for instance, in dealing with perhaps the most widely expressed problem with multifactorial studies of development, as discussed, for instance, by Carl Bereiter. Bereiter argues that frequency-count measures of growth are equivocal, especially when teachers try to apply them to writing instruction, because linguistic devices are interchangeable. The argument is a weighty one. It lies in part behind the current unease with the classic multi-variable studies of writing development by Hunt (1965, 1970), Loban, and Britton. We cannot say that increase in sentence length should be taught, although it correlates positively with age growth, because too many different syntactic constructions may go to produce it, some good but some bad (as a pile-up of adjectives before a noun). The present factoring, however, sorts out variables by clustering them into principal components, thereby reducing the halo of variables one has to consider in dealing with any one variable. Factor 1 here shows t-unit variance, for instance—a maneuver of writing that has traditional and prima-facie approval—associating positively with sentence length and negatively with use of the impersonal "you." That fact confers no preemptive proof. But it does reduce a mass of possible interrelationships to a conceptually modest hypothesis testable with student writers, that rhetorical competence with variety of syntactic length requires some skill in creating long sentences, and that undue reliance upon the "you" pronoun may hold that skill back. Of course, the additional fact that

longer sentences and fewer "you's" characterize older student writing, but not greater t-unit variance, supports that teaching hypothesis.

The ultimate equivocality of intercorrelated measures of text cannot be denied. But against that claim is set the model of writing development sustained by the original study and now by this above factoring. This model envisions writing skill as a complex task, with many interdependencies or layers of "embeddedness" (Lerner), and growth in it as proceeding dialectically, not linearly in lock-step stages but spirally, with much give-and-take, much regression and dormancy, with advent of one skill blocking or even debilitating another (Feldman). The hope has been that with the measurement of enough of the many facets of writing skill, and with application of a certain fund of intuitive experience in teaching and observing students, some of these interactions will become salient and thus arguable. Bereiter's strategy is to launch carefully delimited, context-specific "specialized tasks" (p. 90), a reasonable and valuable tactic. The strategy here has been to launch a full-scale frontal exploration. Here the loss may be in validity, but the gain in food for interpretation.

6. Analysis of writing product

I end with a brief comment on a final problem in interpreting the findings of this study, one that seems to me less of a problem than it is often purported to be. How much does restriction of the investigation to written products, as opposed to writing purposes and processes, weaken an analysis of the findings and curtail their use for teachers?

Certainly analysis of particular findings here often end with a query. All all acts of interpretation do, of course. But that interpretation could carry on longer had research information been gathered about pre-writing intentions, conscious strategies during composing, place and duration of pausing, sequence of revisions or re-scannings, post-writing feelings, and so on. Product alone can only guess at motive, which is so essential for teachers to estimate. Nevertheless, three points need to be clarified about the barrier to interpretation posed by study of product alone. First, so-called process data fetches up against the same barrier. A speak-aloud protocol, a videotape of writing pauses, a pre- or post-writing survey, all

in their own way are products, from which true motives still have to be inferred. Even a writer's assertion of intention is a product, and is not necessary an accurate reflection of actual motivation. With any writer, student or professional, causes of behavior may be unconscious, forgotten, sublimated and otherwise defensed, or merely too complex to sort out.

Granted, the researcher will probably be brought closer to the true motivations by the author's conscious expressions of them than by their manifestation in writing alone. But by the same token, recordings of the writing act or self-reports about writing, alone, will miss some evidence which study of product can provide. Careful analysis of written product especially is useful in describing features of writing which the writers themselves are unaware of. Take passives. Most students cannot report to us much information about their use and misuse of the passive, since they can hardly recognize the form. Older writers can offer more, but we will have difficulty with their testimony in separating their true motives from the advise they have unconsciously picked up from tradition, hearsay, teacher, and textbook. Professional writers who swear they always avoid the passive have been known to use it consistently, sometimes in their sworn affidavits. The developmental study of written product has an advantage in detecting these unconscious or unreported motivations because it describes long-term relationships that writers do not know or think about. According to Hunt (1965), the use of passive in school writing increases from two percent of finite verbs in K4 to nine percent in K12. College students in the present study produced about thirteen percent and the post-graduate writers about fourteen. Despite handbook advise, the rates for professional published writing range from ten to fifteen percent in newspapers and in magazines such as Time, and much higher for scholarly publications. What this developmental data tells us that none of the individual writers could is that increased use of the passive looks like a maturing trait in serious writing (for more on this point, see Haswell 1989b).

Multivariate analysis can go further. Simple correlations often strike fruitful connections unseen by either teachers or writers. For instance, among the four undergraduate groups in the present study, the higher the rate of passive, the lower the rate of sentences with "I" as subject. That suggests a motive for both the use and the avoidance of passive not likely to

be elicited by any spoken protocol. Factor analysis especially tunnels beneath the surface of product traits and offers viable hypotheses for motivation. We have already seen in Factor 8 that passives are associated with production, flow, and cohesion of logical ideas. The implications may counter traditional handbook recommendations, but they make intuitive sense, that the passive is a construction repeatedly useful in continuing old ideas and prompting new ones.

Finally, written product has a certain face value that writing processes lack. The standard appraisal of a writer's worth is not the writer writing but the writer's writing. And that worth is known to the writer, while writing. At every step of the way in the process of writing something is produced, and this something needs to be judged by the writer as product before the process can continue—for instance by scanning, as ironically process studies themselves have discovered. Not only scanning, but every mid-stream act of drafting, revising, and copy-reading involves an appraisal of product. Even that most nebulous point in the act of composing, the moment just before new words are scribed, entails a mental production of text in the head, in various degrees of concrete gestation (Witte). All told, the notion of process as distinct from product approaches nearly the status of canard. At least it is evident that study of written products, however isolated from composing processes, still can provide information that questions, supports, and augments information from other aspects of the social act of writing. The main danger for student, teacher, and research alike is to appraise one aspect without considering the others.

A brief and cautionary conclusion

This last discussion of the value of product-features enters closely into two main conclusions of this presentation of the problems in interpreting the present study and by extension interpreting any study of writing that is empirical, multifactorial, exploratory, and developmental. The first point is that all four of these qualities carry with them both dangers and benefits. Empirical description simplifies the writing studied, yet allows for statistical analysis uncovering trends ordinary scrutiny would not see. The mass

of variables measured precludes a full study of any one variable and hides inter-correlations, yet provides ways to enrich our understanding of a single variable within its true embeddedness and opportunities to sound out inter-correlations. The exploratory nature of the study hides presuppositions because it does a poor job of defining and testing clear-cut hypotheses, yet it is a useful and even necessary method of creating hypotheses in an area actively opening up as an arena of formal study, as is college-age writing development. And any study of development itself risks mistaking nondevelopmental factors for developmental, yet affords the possibility of discovering general psycho-social trends of enormous potential to teaching. The mixture of danger and benefit inheres in all of the interpretive problems discussed above—from the difficulties in conceiving of the true population tapped by the sampling techniques and the boost in applying the findings upon solidifying that conception, to the difficulties in surmising writers' intentions from written product and the insight about unconscious intention toward which such surmise leads. Interpretation of empirical, exploratory findings is a hazard, but a hazard worth the risk.

The second point is that findings, empirical or otherwise, exploratory or otherwise, do not explain themselves. Their interpretation should always take a complex, circular path, from proffered finding, to comparison with findings from intuitive experience and researched studies, to trial and appraisal in praxis, back to new explorations in research. Interpretation is never final or definitive. When it is, it is something else, probably worse.

Works Cited

Bereiter, Carl. "Development in Writing." Cognitive Processes in Writing. Ed. Lee W. Gregg and Erwin Steinberg. Hillsdale NJ: Erlbaum, 1980. 73-93.

Britton, James, et al. The Development of Writing Abilities (11-18). Schools Council Research Studies. London: Macmillan, 1975.

Brossell, Gordon. "Rhetorical Specification in Essay Examination Topics." College English 45 (Feb. 1983): 165-73.

Connors, P. "Research on the Returning or Older Composition Student." ERIC Document Reproduction Service, 1980. ED 192 346.

Cannon, W. W. "Terrors and Affectation: Students' Perceptions of the Writing Process." ERIC Document Reproduction Service, 1981. ED 199 720.

Christensen, Francis, and Bonniejean Christensen. Notes Toward a New Rhetoric: Nine Essays for Teachers. Second Edition. New York: Harper & Row, 1978.

Crowhurst, Marion, and Gene L. Piché. "Audience and Mode of Discourse Effects on Syntactic Complexity in Writing at Two Grade Levels." Research in the Teaching of English 13 (Feb. 1979): 101-9.

Diederich, P. E., S. W. French, and S. T. Carlton. Factors in Judgments of Writing Ability. Research Bulletin RB-61-15. Princeton: Educational Testing Service, 1961 (ERIC Document Reproduction Service, ED 002 172).

Faigley, Lester. "Names in Search of a Concept: Maturity, Fluency, Complexity, and Growth in Written Syntax." College Composition and Communication 31 (Oct. 1980): 291-300.

Faigley, Lester. "The Influence of Generative Rhetoric on the Syntactic Maturity and Writing Effectiveness of College Freshmen." Research in the Teaching of English 13 (Mar. 1979): 197-206.

Faigley, Lester, and Stephen Witte. "Analyzing Revision." College Composition and Communication 32 (1981): 400-14.

Feldman, David Henry. Beyond Universals in Cognitive Development. Norwood, NJ: Ablex, 1980.

Fischer, Kurt W., and Sandra L. Pipp. "Processes of Cognitive Development: Optimal Level and Skill Acquisition." Mechanisms of Cognitive Development, ed. Robert J. Sternberg. New York: W. H. Freedman, 1984. 45-80.

Gaies, S. J. "College Freshman Writing Ability in 1963 and 1977: A Pilot Study Comparison." ERIC Document Reproduction Service, 1980. ED 197 387.

Harman, H. H. Modern Factor Analysis. Second Edition. Chicago: University of Chicago Press, 1967.

Hartvigsen, M. Kip. "A Comparative Study of Quality and Syntactic Maturity Between In-class and Out-of-class Writing Samples of Freshmen at Washington State University." Ph.D. dissertation. Washington State University, 1981.

Haswell, Richard H. "Change in Undergraduate and Post-Graduate Writing Performance: Quantified Findings." ERIC Document Reproduction Service, 1986. ED 269 780.

Haswell, Richard H. "Critique: Length of Text and the Measurement of Cohesion." Research and the Teaching of English 22 (Dec. 1988): 428-33.

Haswell, Richard H. Gaining Ground in College Writing: Tales of Development and Interpretation (1989). Southern Methodist University Press, in press.

Haswell, Richard H. "Textual Research and Coherence: Findings, Intuition, Application." College English 51 (March 1989b): 305-19.

Hilgers, Thomas Lee. "Training College Composition Students in the Use of Freewriting and Problem-solving Heuristics for Rhetorical Invention." Research in the Teaching of English 14 (Dec. 1980): 293-307.

Hunt, Kellogg W. Grammatical Structures Written at Three Grade Levels. NCTE Research Report No. 3. Champaign, IL: National Council of Teachers of English, 1965.

Hunt, Kellogg W. Syntactic Maturity in Schoolchildren and Adults. Monographs of the Society for Research in Child Development 35, no. 1 (1970).

Jewell, Ross M., et al. The Effectiveness of College-level Instruction in Freshman Composition. Final Report. U. S. Office of Education Project No. 2188, Amended. Cedar Falls: University of Northern Iowa, 1969.

Lerner, Richard M. On the Nature of Human Plasticity. Cambridge: Cambridge University Press, 1984.

Linn, Michael D., and Gene Piché. "Black and White Adolescent and Preadolescent Attitudes Toward Black English." Research in the Teaching of English 16 (Feb. 1982): 53-69.

Loban, Walter. Language Development: Kindergarten through Grade Twelve. Urbana, IL: National Council of Teachers of English, 1976.

Mellon, John C. "Issues in the Theory and Practice of Sentence-Combining: A Twenty-year Perspective." Sentence Combining and the Teaching of Writing: Selected Papers from the Miami University Conference. Ed. Donald A. Daiker, et al. Akron OH: University of Akron Press, 1979. 1-38.

Perry, William, Jr. "Cognitive and Ethical Growth: The Making of Meaning." The Modern American College. Eds. Arthur W. Chickering, et al. San Francisco: Jossey-Bass, 1981. 76-116.

Pianko, Sharon. "A Description of the Composing Process of College Freshmen Writers." Research in the Teaching of English 13 (Feb. 1979): 5-22.

Rest, John. Development in Judging Moral Issues. Minneapolis: University of Minnesota Press, 1979.

Schaie, K. W. "A General Model for the Study of Developmental Problems." Psychological Bulletin 64 (1965): 92-107.

Sloan, Gary. "The Subversive Effects of an Oral Culture on Student Writing." College Composition and Communication 30 (May 1979): 156-60.

SPSSX User's Guide. Second Edition. New York: McGraw-Hill 1986.

Swan, M. Beverly. "Sentence Combining in College Composition: Interim Measures and Patterns." Research in the Teaching of English 13 (Dec. 1979): 217-24.

Vygotsky, Lev. S. Mind in Society: The Development of Higher Psychological Processes. Eds. Michael Cole et al. Cambridge MA: Harvard University Press, 1978.

Watson, Cynthia. "Syntactic Change: Writing Development and the Rhetorical Context." The Psychology of Written Language. Ed. Margaret Martlew. London: John Wiley & Sons, 1983. 127-39.

Witte, Stephen. "Pre-text and Composing." College Composition and Communication 38 (Dec. 1987): 397-425.

Zaharias, Jane Ann, and Maia P. Mertz. "Identifying and Validating the Constituents of Literary Response Through a Modification of the Response Preference Measure." Research in the Teaching of English 17 (Oct. 1983): 231-41.